Levy-driven CARMA Random Fields on R?

TERG

Discussion Paper No. 339

Lévy-driven CARMA Random Fields on $\mathbb{R}^n$

Peter J. Brockwell Yasumasa Matsuda

July 2015

TOHOKU ECONOMICS RESEARCH GROUP

GRADUATE SCHOOL OF ECONOMICS AND
MANAGEMENT TOHOKU UNIVERSITY
KAWAUCHI, AOBA-KU, SENDAI,
980-8576 JAPAN

Lévy-driven CARMA Random Fields on $\mathbb{R}^n$

Peter J. Brockwell* Yasumasa Matsuda†

Abstract

We define an isotropic Lévy-driven CARMA(p, q) random field on $\mathbb{R}^n$ as the integral of an isotropic CARMA kernel with respect to a Lévy sheet. Such fields constitute a parametric family characterized by an autoregressive polynomial a and a moving average polynomial b having zeros in both the left and right complex half-planes. They extend the *well-balanced Ornstein-Uhlenbeck process* of Schnurr and Woerner (2011) to a well-balanced CARMA process in one dimension (with a much richer class of autocovariance functions) and to an isotropic CARMA random field on $\mathbb{R}^n$ for $n > 1$. We derive second-order properties of these random fields and find that CAR(1) constitutes a subclass of the well known Matérn class. If the driving Lévy sheet is compound Poisson it is a trivial matter to simulate the corresponding random field on any n-dimensional hypercube. Joint estimation of CARMA kernel parameters and knots locations is proposed in cases driven by compound Poisson sheets and is illustrated by applications to land price data in Tokyo as well as simulated data.

Keywords: compound Poisson, convolution, CARMA random field, Gibbs sampling, knot selection, Lévy sheet, Lévy noise, Matérn class.

1 Introduction

This paper introduces a class of scalar random fields $S_n(t),\ t \in \mathbb{R}^n$, which, when $n = 1$, reduces to a class of non-causal Lévy-driven CARMA processes whose properties have been studied by Brockwell and Lindner (2009).

Traditionally the modelling and analysis of spatial data involves the fitting of a regression model with spatially correlated errors, specified in terms of a parametric family

*Statistics Department, Colorado State University, Fort Collins, Colorado 80523-1877, U.S.A. pjbrock@stat.colostate.edu, Tel: 970 412 3861, Fax: 970 491 7895.

†Graduate School of Economics and Management, Tohoku University, Sendai 980-8576, Japan. matsuda@econ.tohoku.ac.jp

of covariance functions. A large number of families of such covariance functions have been proposed in the literature, in many of which the covariance between the field at any two locations depends only on the Euclidean distance between them. The most frequentlly used family is the Matérn class (Matérn, 1960), for which both the spectral density function and the covariance function can be written explicitly (see e.g. Stein, 1999) . The Matérn class was extended to a class of nonstationary and anisotropic covariances by Paciorek and Schervish (2006).

When fitting covariance models to spatial data, the large sizes of the data sets give rise to significant computational difficulties which mean in particular that maximum likelihood estimation is usually not feasible.

To overcome some of the difficulties Higdon (2002) suggested specifying the field S_n directly as a convolution of a kernel function with a continuous-parameter white noise field. In this paper we utilize this approach, using the white-noise field associated with a second-order Lévy sheet L and a family of kernels g_n which belong to a conveniently parameterized class of kernels reducing, when $n = 1$, to a class of non-causal continuous-time ARMA kernels. The field can be expressed as

$$S_n(t) = \int_{\mathbb{R}^n} g_n(t-u)dL(u), \quad \text{for } t \in \mathbb{R}^n$$

where g_n is specified in Definition 3.1. The resulting covariances between $S_n(t)$ and $S_n(t')$ depend only on the displacement vector $d = t - t'$. Moreover if g_n is isotropic (as it is in this paper) then so is S_n, and the covariances will depend only on the magnitude of d. An especially convenient special case occurs when L is a compound Poisson random measure on the Borel subsets of $\mathbb{R}^n$. In this case S_n can be expressed as

$$S_n(t) := \sum_{i=1}^{\infty} g_n(t - x_i)Y_i,$$

where x_i denotes the location of the i^{th} unit point mass of a Poisson random measure on $\mathbb{R}^n$ and the sequence $\{Y_i\}$ is a sequence of i.i.d. random variables independent of $\{x_i\}$. We shall refer to S_n in this case as a compound Poisson random field with kernel g_n. The random points x_i will be referred to as knots. The knots contained in any bounded subset $D \subset \mathbb{R}^n$, given their number, are uniformly and independently distributed on D.

As pointed out by Higdon (2002), the convolution model permits dimension reduction in the following sense. If we consider only those contributions to the field from a finite number, M, of knots contained in a bounded subset of $\mathbb{R}^n$, then the necessary computations for estimation and kriging can all be carried out in terms of $M \times M$ matrices rather than the typically much larger $T \times T$ matrices, where T is the number of observations. Moreover the convolution approach can be extended readily to deal with nonstationary,

anisotropic, multivariate and spatio-temporal modeling (see e.g. Fuentes, 2002; Calder and Cressie, 2007; Majumdar et al, 2010; Sampson, 2010).

In the convolution approach the choice of kernels and knot locations plays a critical role for estimation and kriging performance. Higdon (2002) introduced a family of kernels that gave standard classes of Gaussian, exponential and spherical covariances, and suggested ad hoc choices for them in his ozone modeling examples. In most existing studies knot locations are either regularly or randomly spaced points, the number being chosen so as not to exceed computational limits. For given knot locations, the kernel parameters can be estimated by either classical maximum likelihood or Bayesian methods, the specification of knot locations and the estimation of kernel parameters being conducted separately.

In this paper we determine the first and second-order properties of the scalar random fields defined above. The family of isotropic CARMA(p, q) kernels specified in Definition 3.1 generates a rich new class of spatial covariances which are not necessarily non-negative and not necessarily monotonically decreasing. For model-fitting and kriging we use a Bayesian Markov chain Monte-Carlo method to estimate the knot locations and CARMA parameters simultaneously. We find in our examples that joint estimation of the kernel parameters and knot locations improves substantially on estimation and kriging in which the knot locations are specified prior to kernel estimation. For the CAR(1) field it is found that the covariance function belongs to the Matérn class with smoothness parameter $\frac{n}{2} + 1$. For the general CARMA(p, q) field the spectral density is found explicitly and the covariance function is expressed as a modified Hankel transform (a one-dimensional integral) of an explicitly defined function. It is evaluated explicitly for $n = 1$ and $n = 3$. If the Lévy noise is derived from a compound Poisson sheet then it is a trivial matter to simulate a corresponding CARMA random field without any matrix operations. We make use of this in our simulated examples. We observe also that if the CARMA kernel is replaced by a Matérn kernel with smoothness parameter ν then the resulting field has a Matérn covariance function with smoothness parameter $\nu + \frac{n}{2}$.

The stationary and isotropic CARMA random fields introduced in this paper suggest further studies, in which we plan to extend the models to represent nonstationary and anisotropic random fields. The paper is organized as follows. Section 2 introduces some background tools relating to Lévy sheets and the multivariate Fourier transform of an isotropic function. In Section 3 we evaluate the first and second-order moments of a CARMA random field driven by a second-order Lévy sheet. Section 4 discusses the joint estimation of CARMA parameters and knot locations. Section 5 demonstrates the performance of CARMA random fields fitted to both simulated data and to a real-life example involving Tokyo land price data.

2 Preliminaries

Lévy noise and Lévy sheets have recenty received much attention in connection with stochastic partial differential equations (see e.g. Koshnevisan and Nualart (2008)). In this paper we make use of them to define multiparameter versions of single-parameter CARMA (continuous-time autoregressive moving-average) processes.

If L is a scalar n-parameter Lévy sheet and $\dot{L}$ is the corresponding Lévy noise as defined by Koshnevisan and Nualart (2008)), Section 2, then we can follow their construction to define an n-parameter scalar CARMA random field via the integral

$$S_n(t) = \int_{\mathbb{R}^n} g_n(t-u) dL(u),\ t \in \mathbb{R}^n, \tag{1}$$

where the kernel $g_n : \mathbb{R}^n \to \mathbb{R}$ is a suitably chosen analogue of the one-parameter CARMA kernel. Kernels of the form $g_n(t) = \phi(\|t\|)$ and $g_n(t) = \prod_{i=1}^n \psi_i(t_i)$ are of particular interest but in this paper we shall restrict attention to the former and refer to the corresponding random field as *isotropic*.

Important cases arise (i) when L is a Brownian sheet, in which case $\dot{L}$ is Gaussian white noise and (ii) when L is a compound Poisson sheet, in which case L is the random measure on the Borel subsets of $\mathbb{R}^n$ given by,

$$L(A) = \sum_{i=1}^{\infty} Y_i \mathbf{1}_{x_i}(A),\ A \in \mathcal{B}(\mathbb{R}^n), \tag{2}$$

where x_i denotes the location of the i^{th} unit point mass of a Poisson random measure on $\mathbb{R}^n$ and the sequence $\{Y_i\}$ is a sequence of i.i.d. random variables independent of $\{x_i\}$.

A variety of fields are generated by (1) depending on the choice of L and of the kernel g_n. Our prime concern will be with second order fields and their first and second order properties which we shall apply to the modelling of multiparameter data sets. If L is a second-order Lévy sheet then there exist constants $\mu \in \mathbb{R}$ and $\sigma > 0$ such that for all real-valued functions g and h in $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$,

$$\mathbb{E} \int_{\mathbb{R}^n} g(t)\ dL(t) = \mu \int_{\mathbb{R}^n} g(t)\ dt \tag{3}$$

and

$$\text{Cov}\left[\int_{\mathbb{R}^n} g(t)\ dL(t), \int_{\mathbb{R}^n} h(t)\ dL(t)\right] = \sigma^2 \int_{\mathbb{R}^n} g(t)h(t)\ dt. \tag{4}$$

If the kernel g_n in definition (1) belongs to $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ then the first and second-order properties of the field S_n are given by

$$\mathbb{E}S_n(t) = \mu \int_{\mathbb{R}^n} g_n(u) du \tag{5}$$

and

$$\gamma_n(t) := \mathrm{Cov}(S_n(t), S_n(0)) = \sigma^2 g_n * g_n(t),\ t \in \mathbb{R}^n, \tag{6}$$

i.e. σ^2 times the multi-dimensional convolution of the kernel g_n with itself.

In order to compute the autocovariance function (6) we shall first compute the multivariate Fourier transform $\tilde{g}_n = F_n f$ of the kernel g_n, defined, for $g_n \in L^1(\mathbb{R}^n)$, by

$$\tilde{g}_n(\omega) = F_n g_n(\omega) := (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-i\langle \omega, t\rangle} g_n(t)\ dt,\ \omega \in \mathbb{R}^n. \tag{7}$$

Then by (6), provided $(\tilde{g}_n)^2 \in L^1(\mathbb{R}^n)$, the autocovariance function γ_n is $(2\pi)^{n/2}$ times the inverse Fourier transform of $\sigma^2(\tilde{g}_n)^2$, i.e.

$$\gamma_n(t) = \sigma^2 \int_{\mathbb{R}^n} e^{i\langle \omega, t\rangle} \tilde{g}_n(\omega)^2\ d\omega,\ t \in \mathbb{R}^n, \tag{8}$$

so that the multivariate spectral density function f_n of S_n is

$$f_n(\omega) = \sigma^2 \tilde{g}_n(\omega)^2,\ \omega \in \mathbb{R}^n. \tag{9}$$

If the function g_n is radial, i.e. if there is a univariate function ϕ such that

$$g_n(t) = \phi_n(\|t\|),\ t \in \mathbb{R}^n, \tag{10}$$

then it is well-known (see e.g. Nowak and Stempak (2014)) that the Fourier transform of g_n is also radial and can be expressed as

$$\tilde{g}_n(\omega) = \tilde{\phi}_n(\|\omega\|) := H_{\frac{n}{2}-1}\phi_n(\|\omega\|),\ \omega \in \mathbb{R}^n, \tag{11}$$

where H_m denotes the modified Hankel transform,

$$H_m g(x) = \int_0^\infty g(y) \frac{J_m(xy)}{(xy)^m} y^{2m+1}\ dy,\ x > 0,\ m \geq -1/2, \tag{12}$$

and J_m is the Bessel function of the first kind of order m. Clearly the function $\tilde{g}_n^2$ is also radial so that (8) can be reexpressed as

$$\gamma_n(t) = \sigma^2 (2\pi)^{n/2} H_{\frac{n}{2}-1}\tilde{\phi}_n^2(\|t\|),\ t \in \mathbb{R}^n. \tag{13}$$

If, for each $n \in \mathbb{N}$, g_n is a radial function in $L^1(\mathbb{R}^n)$ with $g_n(t) = \phi(\|t\|),\ t \in \mathbb{R}^n$, where $\phi : \mathbb{R} \to \mathbb{R}$ is a function independent of n, then we can write the Fourier transform (7) as

$$\tilde{g}_n(\omega) = F_n g_n(\omega) = \mathcal{F}_n \phi(\|\omega\|),\ n \in \mathbb{N}, \tag{14}$$

where, by a very elegant result of Grafakos and Teschl (2013), the mapping $\mathcal{F}_{n+2}$ is related to the mapping $\mathcal{F}_n$ by

$$\mathcal{F}_{n+2}\phi(r) = -\frac{1}{r}\frac{d}{dr}\mathcal{F}_n \phi(r),\ n \in \mathbb{N}. \tag{15}$$

Equation (15) differs by a factor 2π from equation (2) of Grafakos and Teschl owing to their slightly different definition of the Fourier transform. Thanks to this equation it becomes a simple matter to compute $\mathcal{F}_n\phi, n \in \mathbb{N}$, from $\mathcal{F}_1\phi$ and $\mathcal{F}_2\phi$.

We shall say that the field S_n defined by (1) is *isotropic* if the kernel g_n satisfies (10) for some function $\phi : \mathbb{R} \to \mathbb{R}$. (The function ϕ may depend on n but in all of our applications it will denote a function independent of n.)

3 Isotropic Lévy-driven CARMA random fields

Definition 3.1. *Let $a_*(z) = z^p + a_1 z^{p-1} + \cdots a_p = \prod_{i=1}^p (z - \lambda_i)$ be a polynomial of degree p with real coefficients and distinct zeroes $\lambda_1, \ldots, \lambda_p$ having strictly negative real parts and let $b_*(z) = b_0 + b_1 z + \cdots b_q z^q = \prod_{i=1}^q (z - \xi_i)$ with real coefficient b_j and $0 \leq q < p$. Suppose also that $\lambda_i \neq \mu_j$ for all i and j. Then defining*

$$a(z) = \prod_{i=1}^{p} (z^2 - \lambda_i^2) \quad \text{and} \quad b(z) = \prod_{i=1}^{q} (z^2 - \xi_i^2),$$

the isotropic L-driven CARMA(p, q) field with autoregressive polynomial a and moving average polynomial b is

$$S_n(t) = \int_{\mathbb{R}^n} \sum_{r=1}^{p} \frac{b(\lambda_r)}{a'(\lambda_r)} e^{\lambda_r \|t-u\|} dL(u), \ t \in \mathbb{R}^n, \tag{16}$$

where a' denotes the derivative of the polynomial a and $\|t - u\|$ denotes the Euclidean norm of the vector $t - u$. □

Remark 3.2. The kernel $g(t) = \sum_{i=1}^p b(\lambda_i) e^{\lambda_i \|t\|} / a'(\lambda_i)$ appearing in (16) is the special case, when $a(z)$ has distinct zeroes, of the kernel,

$$g(t) = \sum_{\lambda : \mathcal{R}(\lambda) < 0} \mathbf{Res}_{z=\lambda} \left[e^{z\|t\|} b(z)/a(z) \right],$$

where the sum is over the *distinct* zeroes of $a(z)$ with negative real parts and $\mathbf{Res}_{z=\lambda} f(z)$ denotes the residue at $z = \lambda$ of the function f. When $n = 1$, it is the kernel of the non-causal CARMA$(2p, 2q)$ process with autoregressive and moving average polynomials $a(z)$ and $b(z)$ respectively (see Brockwell and Lindner (2009)). For simplicity of exposition and with no essential loss of generality we shall assume throughout this paper that the zeroes of $a(z)$ are distinct. In the special case when $a(z) = z^2 - \lambda^2, \ \lambda < 0$, and $b(z) = 1$, the process S_1 defined by (16) reduces to $(2\lambda)^{-1}$ times the *well-balanced Ornstein-Uhlenbeck process* of Schnurr and Woerner (2011). The definition (16) can therefore be regarded in their terminology as a *well-balanced CARMA process* when $n = 1$ and as the generalization of such a process to a field on $\mathbb{R}^n$ when $n > 1$. .

For the compound Poisson L defined by (2),

$$S_n(t) := \sum_{i=1}^{\infty} \sum_{r=1}^{p} \frac{b(\lambda_r)}{a'(\lambda_r)} Y_i e^{\lambda_r ||t-x_i||}, \ t \in \mathbb{R}^n, \tag{17}$$

an expression which permits very simple simulation of S_n on any hypercube or on certain more complicated bounded subsets of $\mathbb{R}^n$.

Remark 3.3. It is interesting to note that the kernel

$$\sum_{r=1}^{p} \frac{b(\lambda_r)}{a'(\lambda_r)} e^{\lambda_r \|t\|} = \gamma(\|t\|),$$

where γ is the autocovariance function of the one-dimensional CARMA(p, q) process with autoregressive polynomial $a_*(z)$,moving average polynomial $b_*(z)$ and driving Lévy process L such that $Var(L_1) = 1$. This result can be derived from the representation (65) of γ in Brockwell (2014).

Our first theorem establishes the first and second-order properties of the isotropic CAR(1) field with autoregressive polynomial $a(z) = z^2 - \lambda^2, \lambda < 0$.

Theorem 3.4. *If L is a second-order Lévy sheet satisfying (3) and (4) and if the defining polynomials in (16) are $a(z) = z^2 - \lambda^2$ (where $\lambda < 0$) and $b(z) = 1$, we shall refer to S_n as an isotropic CAR(1) field with kernel $g_n(t) = (2\lambda)^{-1} e^{\lambda \|t\|} = \phi(\|t\|), \ t \in \mathbb{R}^n$, where $\phi(r) := e^{\lambda r}/(2\lambda), \ r \geq 0$. Then*

$$\mathbb{E}[S_n(t)] = \frac{\mu \pi^{\frac{n}{2}} \Gamma(n+1)}{2\lambda |\lambda|^n \Gamma(\frac{n}{2}+1)}, \tag{18}$$

the spectral density of S_n is

$$f_n(\omega) = \sigma^2 c_n^2 (\|\omega\|^2 + \lambda^2)^{-n-1}, \ \omega \in \mathbb{R}^n, \tag{19}$$

where

$$c_n = \begin{cases} -2^{n/2-1} \Gamma(\frac{n+1}{2})/\sqrt{\pi} & \text{if } n \text{ is odd,} \\ -2^{-n/2} \Gamma(n)/\Gamma(\frac{n}{2}) & \text{if } n \text{ is even,} \end{cases}$$

and the autocovariance function of S_n is

$$\gamma_n(t) = \sigma^2 c_n^2 \left(\frac{\pi}{2}\right)^{\frac{n}{2}} \frac{\|\lambda t\|^{\frac{n}{2}+1}}{|\lambda|^{n+2} \Gamma(n+1)} K_{\frac{n}{2}+1}(\|\lambda t\|), \ t \in \mathbb{R}^n \tag{20}$$

where $K_{\frac{n}{2}+1}$ denotes the modified Bessel function of the second kind of order $\frac{n}{2} + 1$.

Proof. From (5) we find at once that $\mathbb{E}[S_n(t)] = \frac{\mu}{2\lambda}\int_{R^n} e^{\lambda\|t\|}dt$. Rewriting the integral in polar coordinates then gives

$$\mathbb{E}[S_n(t)] = \frac{\mu}{2\lambda}\int_0^\infty e^{\lambda r}\frac{nr^{n-1}\pi^{n/2}}{\Gamma(\frac{n}{2}+1)}\,dr = \frac{\mu\pi^{\frac{n}{2}}\Gamma(n+1)}{2\lambda|\lambda|^n\Gamma(\frac{n}{2}+1)}.$$

In order to compute the spectral density and autocovariance function of S_n we shall first determine $\tilde{g}_n(\omega) = F_n g_n(\omega) = \mathcal{F}_n\phi(\|\omega\|) = \tilde{\phi}_n(\|\omega\|),\ \omega \in \mathbb{R}^n$, for $n = 1$ and 2, and then use these, in conjunction with (15), to determine $\tilde{g}_n,\ n \in \mathbb{N}$. The spectral density of S_n will then be found from (9) and the autocovariance function from (13).
The case $n = 1$:

Substituting $\phi(r) = e^{\lambda r}/(2\lambda)$ in (7) and evaluating the integral gives

$$\tilde{g}_1(\omega) = -\frac{1}{\sqrt{2\pi}(\omega^2+\lambda^2)},\ \omega \in \mathbb{R}. \tag{21}$$

The case $n = 2$: From (11) and (12) we have

$$\tilde{g}_2(\omega) = \frac{1}{2\lambda}H_0 e^{\lambda(\cdot)}(\|\omega\|) = \frac{1}{2\lambda}\int_0^\infty e^{\lambda y}J_0(\|\omega\|\,y)y\,dy \;=\; -\frac{1}{2(\|\omega\|^2+\lambda^2)^{3/2}},\quad \omega \in \mathbb{R}^2, \tag{22}$$

where the last equality can be found from tables of the zero-order Hankel transform (see e.g. Piessens (2010), table 9.1, equation(4)).

From (21) and (22) it follows that $\mathcal{F}_1\phi(r) = -\frac{1}{\sqrt{2\pi}}(r^2+\lambda^2)^{-1}$ and $\mathcal{F}_2\phi(r) = -\frac{1}{2}(r^2+\lambda^2)^{-3/2}$. Successive application of (15) with $n = 1, 2, 3, \ldots$, then gives,

$$\mathcal{F}_n\phi(r) = \tilde{\phi}_n(r) = c_n(r^2+\lambda^2)^{-\frac{n+1}{2}},\ n \in \mathbb{N}, \tag{23}$$

with c_n defined as in (19). These equations, with (9) and (14) immediately give equation (19) for the spectral density of S_n.

Substituting from (23) into equation (13) and using Piessens (2010), Table 9.2, equation (12), we find immediately that the autocovariance function of S_n is given by (20) as claimed. □

Remark 3.5. The values of $K_r(z), r \geq 0, z \in \mathbb{C}$, are available in Matlab as besselk(r, z). The functions $K_{n+\frac{1}{2}},\ n \in \mathbb{N}$, take a particularly simple form. From Abramowitz and Stegun (2012), equations (10.2.17), we have

$$K_{\frac{3}{2}}(x) = \sqrt{\pi/(2z)}e^{-z}(1+z^{-1}) \tag{24}$$

and

$$K_{\frac{5}{2}}(x) = \sqrt{\pi/(2z)}e^{-z}(1+3z^{-1}+3z^{-2}). \tag{25}$$

Moreover, if we define

$$h_n(z) := (-1)^{n+1}\sqrt{\pi/(2z)}K_{n+\frac{1}{2}(z)},\ n \in \mathbb{N},$$

we can readily calculate $K_{n+\frac{1}{2}}(z)$, for integers $n > 2$ from (24), (25) and the recursions (Abramowitz and Stegun (2012), equation (10.2.18)),

$$h_{n+1}(z) = h_{n-1}(z) - (2n+1)z^{-1}h_n(z).$$

Remark 3.6. Direct application of Theorem 3.4 gives the following results for the important cases $n = 1, 2$, and 3.

$$\mathbb{E}[S_1] = -\frac{\mu}{\lambda^2}, \quad f_1(\omega) = \frac{\sigma^2}{2\pi(\omega^2+\lambda^2)^2}, \qquad \gamma_1(t) = \frac{\sigma^2}{4|\lambda|^3}e^{-|\lambda t|}(1+|\lambda t|).$$

$$\mathbb{E}[S_2] = \frac{\pi\mu}{\lambda^3}, \qquad f_2(\omega) = \frac{\sigma^2}{4(\|\omega\|^2+\lambda^2)^3}, \qquad \gamma_2(t) = \frac{\sigma^2\pi}{16\lambda^4}\|\lambda t\|^2 K_2(\|\lambda t\|),$$

$$\mathbb{E}[S_3] = -\frac{4\pi\mu}{\lambda^4},\ f_3(\omega) = \frac{2\sigma^2}{\pi(\|\omega\|^2+\lambda^2)^4}, \quad \gamma_3(t) = \frac{\sigma^2\pi}{12|\lambda|^5}e^{-\|\lambda t\|}(3+3\|\lambda t\|+\|\lambda t\|^2).$$

Remark 3.7. The general Matérn autocovariance function has the form (see e.g. Stein (1999))

$$\gamma_n(t) = \sigma^2(\|at\|)^\nu K_\nu(\|at\|),\ a > 0,\ \nu > 0.$$

If we define S_n as in (1) with a Matérn *kernel*, i.e. with

$$g_n(t) = (\|at\|)^\nu K_\nu(\|at\|),\ \nu > 0, \tag{26}$$

then we find, by arguments analogous to those used in the proof of Theorem 3.4, that the corresponding mean, spectral density and autocovariance function of S_n are,

$$\mathbb{E}[S_n] = \mu\ 2^{\nu-1}\left(\frac{4\pi}{a^2}\right)^{\frac{n}{2}}\Gamma(\frac{n}{2}+\nu), \qquad f_n(\omega) = \frac{\sigma^2(2a^2)^{2\nu}2^{n-2}\Gamma\left(\frac{n}{2}+\nu\right)^2}{(\|\omega\|^2+a^2)^{n+2\nu}} \quad \text{and}$$

$$\gamma_n(t) = \frac{\sigma^2}{2}\left(\frac{2\pi}{a^2}\right)^{n/2}\frac{\Gamma\left(\frac{n}{2}+\nu\right)^2}{\Gamma(n+2\nu)}\ (\|at\|)^{\frac{n}{2}+2\nu}K_{\frac{n}{2}+2\nu}(\|at\|).$$

In other words the Matérn kernel with index ν generates a sheet with Matérn autocovariance function having index $\frac{n}{2}+\nu$. The results of Theorem 3.4 can be derived from these more general results since

$$(2\lambda)^{-1}e^{\lambda\|t\|} = (2\lambda)^{-1}e^{-\|\lambda t\|} = \frac{1}{\lambda\sqrt{2\pi}}\|\lambda t\|^{\frac{1}{2}}K_{\frac{1}{2}}(\|\lambda t\|).$$

We next consider the isotropic CARMA(p,q) field defined by (16).

Theorem 3.8. *If L is a second-order Lévy sheet satisfying (3) and (4) and if the isotropic CARMA field S_n is specified as in Definition 3.1, i.e. with kernel $g_n(t) = \sum_{r=1}^{p} \frac{b(\lambda_r)}{a'(\lambda_r)} e^{\lambda_r \|t\|}$, $t \in \mathbb{R}^n$, then*

$$\mathbb{E}[S_n(t)] = \mu \sum_{i=1}^{p} \frac{b(\lambda_i)}{a'(\lambda_i)} \frac{\pi^{\frac{n}{2}} \Gamma(n+1)}{|\lambda|^n \Gamma(\frac{n}{2}+1)}, \tag{27}$$

and the spectral density of S_n is

$$f_n(\omega) = \rho_n(\|\omega\|) = \begin{cases} \sigma^2 \frac{2^{n-2}}{\pi} \left(\psi_1^{(\frac{n-1}{2})}(r)\right)^2, & \text{if } n \text{ is odd,} \\ \sigma^2 2^{n-3} \left(\psi_2^{(\frac{n-2}{2})}(r)\right)^2, & \text{if } n \text{ is even,} \end{cases} \tag{28}$$

where $r = -\|\omega\|^2$, the superscripts on the functions ψ_1 and ψ_2 denote order of differentiation with respect to r, and

$$\psi_k(r) = -\sum_{i=1}^{p} \frac{2\lambda_i b(\lambda_i)}{a'(\lambda_i)(-r+\lambda_i^2)^{\frac{k+1}{2}}}, \quad k = 1, 2. \tag{29}$$

The autocovariance function of S_n is

$$\gamma_n(t) = (2\pi)^{n/2} H_{\frac{n}{2}-1} \rho_n(\|t\|), \tag{30}$$

where H_m denotes the modified Hankel transform of order m as defined in (12) and ρ_n is defined as in (28).

Proof. For the field S_n with kernel $g_n(t) = \sum_{r=1}^{p} \frac{b(\lambda_r)}{a'(\lambda_r)} e^{\lambda_r \|t\|}$, $t \in \mathbb{R}^n$, equation (5) gives

$$\mathbb{E}S_n(t) = \mu \sum_{i=1}^{p} \frac{b(\lambda_i)}{a'(\lambda_i)} \int_{\mathbb{R}^n} e^{\lambda_i \|u\|} du.$$

The integral on the right was evaluated in the proof of Theorem 3.4. Referring to equation (18) we see at once that (27) holds.

The Fourier transform of $e^{\lambda_i \|t\|}$, $t \in \mathbb{R}^n$, is, from (23), $2\lambda_i c_n (\|\omega\|^2 + \lambda_i^2)^{\frac{n+1}{2}}$, $\omega \in \mathbb{R}^n$, where the coefficients c_n are as in (20). The Fourier transform of g_n is therefore

$$\tilde{g}_n(\omega) = c_n \sum_{i=1}^{p} \frac{2\lambda_i b(\lambda_i)}{a'(\lambda_i)(\|\omega\|^2 + \lambda_i^2)^{\frac{n+1}{2}}}.$$

With ψ_1 and ψ_2 defined as in (29) it is a straightforward calculation to show that

$$\tilde{g}_n(\omega) = \begin{cases} \frac{2^{(n-2)/2}}{\sqrt{\pi}} \psi_1^{(\frac{n-1}{2})}(r), & \text{if } n \text{ is odd,} \\ \\ 2^{(n-3)/2} \psi_2^{(\frac{n-2}{2})}(r), & \text{if } n \text{ is even,} \end{cases}$$

where $r = -\|\omega\|^2$. The spectral density of S_n is then $f_n(\omega) = \sigma^2 \tilde{g}_n(\omega)^2$, i.e. (28).

The autocovariance function of S_n is $(2\pi)^{n/2}$ times the inverse Fourier transform of the spectral density and, because the spectral density is radial, it can be expressed as the one-dimensional integral (30). □

A fast algorithm for the numerical computation of Hankel transforms was developed by Knockaert (2000). If the dimensionality n is odd it is possible to express the spectral density of S_n in a simpler form and to use this to compute the autocovariance functions explicitly in the particular cases when $n = 1$ and $n = 3$.

Corollary 3.9. *The function ψ_1 in (28) simplifies to*

$$\psi_1(r) = \frac{\beta(r)}{\alpha(r)}, \tag{31}$$

where $\beta(z) := \prod_{i=1}^{q}(z - \xi_i^2)$ and $\alpha(z) := \prod_{i=1}^{p}(z - \lambda_i^2)$. This implies, by (28) that when n is odd the spectral density $f_n(\omega)$ of S_n is a rational function of $\|\omega\|^2$, easily calculated from (28).

Corollary 3.10. *In the special cases $n = 1$ and $n = 3$ we have*

$$\gamma_1(t) = \sigma^2 \sum_{i=1}^{p} \mathbf{Res}_{z=\lambda_i} \left[e^{z|t|} \frac{b(z)^2}{a(z)^2} \right]$$

and

$$\gamma_3(t) = -\sigma^2 \frac{2\pi}{\|t\|} \sum_{i=1}^{p} \mathbf{Res}_{z=\lambda_i} \left[e^{z\|t\|} \frac{(a'(z)b(z) - a(z)b'(z))^2}{z a(z)^4} \right].$$

Proof. From (28) and (31) with $n = 1$,

$$f_1(\omega) = \frac{\sigma^2}{2\pi} \left(\frac{\beta(-\omega^2)}{\alpha(-\omega^2)} \right)^2.$$

Hence

$$\gamma_1(t) = \frac{\sigma^2}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} \left(\frac{\beta(-\omega^2)}{\alpha(-\omega^2)} \right)^2 d\omega.$$

Changing the variable of integration to $z = i\omega$, using the relations $\alpha(z^2) = a(z)$ and $\beta(z^2) = b(z)$ and using contour integration to evaluate the integral gives the required expression for γ_1.

From (28) and (31) with $n = 3$,

$$f_3(\omega) = \frac{2\sigma^2}{\pi} \delta(-\|\omega\|^2),$$

where

$$\delta(r) = \alpha(r)^{-4}(\alpha(r)\beta'(r) - \alpha'(r)\beta(r))^2.$$

Hence

$$\gamma_3(t) = \frac{2\sigma^2}{\pi} \int_{\mathbb{R}^3} e^{i\langle \omega, t\rangle} \delta(-\|\omega\|^2) d\omega.$$

Expressing this integral in three-dimensional polar coordinates and integrating out the angular components gives

$$\gamma_3(t) = \frac{4\sigma^2}{\|t\|} \mathbf{Im} \int_{-\infty}^{\infty} \rho \delta(-\rho^2)^2 e^{i\rho\|t\|} d\rho.$$

Rewriting δ in terms of the polynomials a and b and again using contour integration to evaluate the resulting integral gives the required expression for γ_3. □

Example 3.11. *A two-parameter CAR(2) random field.*
If L is a two-dimensional second-order Lévy sheet satisfying (3) and (4) and $S_2(t), t \in \mathbb{R}^2$, is the two-parameter random field defined by (16) with $a(z) = (z^2 - \lambda_1^2)(z^2 - \lambda_2^2)$, $\lambda_1 = -1 + 2i$, $\lambda_2 = -1 - 2i$ and $b(z) = 1$, then the kernel g_2 is readily calculated as

$$g_2(t) = \phi_2(\|t\|) = \frac{e^{-\|t\|}}{20} \left(\cos(2\|t\|) + \frac{1}{2} \sin(2\|t\|) \right).$$

From (28) and (29) we then find that the spectral density of S_2 is

$$f_2(\omega) = \rho_2(\|\omega\|) = \frac{\sigma^2}{2(\lambda_1^2 - \lambda_2^2)^2} \left[(\|\omega\|^2 + \lambda_2^2)^{-3/2} - (\|\omega\|^2 + \lambda_2^2)^{-3/2} \right]^2$$

and, from (30) and (12),

$$\gamma_2(t) = 2\pi H_0 \rho_2(\|t\|) = 2\pi \int_0^{\infty} y \rho_2(y) J_0(ty)\, dy.$$

The latter integral is easily evaluated using the Matlab function besselj to compute $J_0(ty)$. Numerical evaluations of correlation functions of two-parameter CAR(1) and CAR(2) random fields will be demonstrated in Sec. 5.

4 Estimation and kriging

4.1 Parameter estimation

This section considers estimation of the parameters of the CARMA kernel when the driving Lévy sheet is compound Poisson as in (17). The data consist of observations of $S_n(s)$ at the finite number of sampling points $s_1, \ldots, s_T$, scattered irregularly over a bounded region $D \subset \mathbb{R}^n$.

Our model for the data will be a modified form of the model (17) which allows in particular for measurement errors. We shall write the CARMA kernel, with parameter vector $\theta = (\lambda_1, \ldots, \lambda_p)'$ as

$$g_\theta(s,x) = \sum_{r=1}^{p} \frac{b(\lambda_r)}{a'(\lambda_r)} e^{\lambda_r ||s-x||}.$$

Our model is then,

$$\begin{aligned} S_n(s_j) &= Z_j\beta + \sum_{i=1}^{M} Y_i g_\theta(s_j, x_i) + \varepsilon_j, \;\; j = 1, \ldots, T, \\ Y_i &\sim iidN(0,\tau^2), \varepsilon_j \sim iidN(0,\delta^2). \end{aligned} \tag{32}$$

This model differs from (17) in several respects. First, the mean is accounted for by the term $Z_j\beta$, corresponding to the regressor vector $Z = (Z_1', \ldots, Z_T')'$. Secondly, independent normally distributed noise terms ε_j have been included to represent measurement errors, known in the literature as nugget effects. Finally, the 'knot points', x_i, are the points of the underlying spatial Poisson process falling in the region D and their number, $n(D)$, has the Poisson distribution. Conditionally on the value of $n(D)$, the knot points are independently and uniformly distributed on D. The number of knots and their locations are crucial in the estimation procedure as discussed in section 4.2.

The model (32) resembles a standard regression model with the values of the kernel function $g_\theta(s, x_i)$ as independent variables. There are however difficulties which are not encountered in classical regression analysis. (i) The knots x_i, and hence the independent variables $g_\theta(s, x_i)$, are stochastic. (ii) The independent variables depend on the parameter θ to be estimated. (iii) The regression coefficients Y_i are not fixed but random. These difficulties prohibit classical maximum likelihood estimation and so a different approach is needed.

We start by evaluating the conditional likelihood function of the parameters $(\beta, \delta^2, \theta, \tau^2)$ given the number, M, and locations, $\{x_i\}$, of the knots in the model (32). The conditional variance of $S_n = (S_n(s_1), \ldots, S_n(s_T))'$ is then

$$\tau^2 V_\theta V_\theta' + \delta^2 I_T = \delta^2 \left(I_T + \kappa^2 V_\theta V_\theta' \right) = \delta^2 R,$$

say, where V_θ is the T by M matrix whose $(p,q)^{\text{th}}$ component is $V_{pq} = g_\theta(s_p, x_q)$ and $\kappa^2 = \tau^2/\delta^2$. The conditional likelihood is

$$S_n \sim N\left(Z\beta, \delta^2 R\right).$$

Calculation of the inverse and determinant of R requires an infeasible $O(T^3)$ operations. We apply Corollaries 18.1.2 and 18.2.10 of Harville (1997) to calculate them in $O(M^3)$

operations as

$$(\delta^2 R)^{-1} = \delta^{-2} \left\{ I_T - V_\theta (\kappa^{-2} I_M + V_\theta' V_\theta)^{-1} V_\theta' \right\},$$

and

$$\left| \delta^2 R \right| = \delta^{2T} \left| \kappa^2 V_\theta' V_\theta + I_M \right|.$$

These expressions yield, in $O(M^3)$ operations, the conditional log likelihood of S_n (a function (of $\Omega = (\beta, \delta^2, \theta, \kappa^2, \{x_i\})$) in $O(M^3)$ operations as

$$\log L(S_n|\Omega) = -\frac{T}{2} \log \delta^2 - \frac{1}{2} \log |R| - \frac{1}{2\delta^2} (S_n - Z\beta)' R^{-1} (S_n - Z\beta), \tag{33}$$

It is infeasible to estimate $(\beta, \delta^2, \theta, \kappa^2)$ by maximizing (33) unless $x_1, \ldots, x_M$, are fixed and known, which leads us to take a Bayesian approach.

For given M, the knot locations, $\{x_i\}$, are independently and uniformly distributed on D. Taking this as the prior distribution for $\{x_i\}$ opens the way to Bayesian estimation of the model. Bayesian inference requires us to specify priors for the other model parameters. The Gibbs sampler is used to draw samples of parameters from the following posterior

$$p(\Omega|S_n) \propto L(S_n|\Omega) p(\beta) p(\delta^2) p(\theta) p(\kappa^2) p(\{x_i\}).$$

We assign the conjugate priors $N(\beta_0, \Sigma_0)$ and $Ga(a_1, a_2)$ for β and δ^{-2}, respectively. We update β from $N(\mu_\beta, \Sigma_\beta)$, where

$$\Sigma_\beta^{-1} = \Sigma_0^{-1} + \delta^{-2} Z' R^{-1} Z,$$

and

$$\mu_\beta = \Sigma_\beta \left(\delta^{-2} Z' R^{-1} S_n + \Sigma_0^{-1} \beta_0, \right)$$

and update δ^{-2} from the gamma distribution with parameters $a_1 + T/2$ and $a_2 + 1/2(S_n - Z\beta)' R^{-1} (S_n - Z\beta)$. For the other parameters without conjugate prior distributions, we draw samples using the Metropolis-Hastings (MH) algorithm. We employ the vague prior distribution, uniform on the parameter set for (θ, κ^2), and use random walk chains driven by normal variables with mean 0 to update them in the MH algorithm. Priors for the knot locations and the updating procedure will be discussed in detail in the following section.

4.2 Knot specification

Specification of the knot locations $\{x_i\}$ plays a critical role in estimation and kriging using CARMA models. Gelfand et al. (2012) proposed a way in the predictive process approach

to specify them separately from parameter estimation. Zhang et al. (2015) employed a procedure for joint estimation for knots and parameter in the context of composite approach. We shall follow the latter approach with a modification to allow for our assumed uniform distribution of knot locations.

Suppose we have a set $U = \{u_1, \ldots, u_N\} \subset D$ of knot candidates, which are independent uniformly distributed points in D. We take the prior distribution of the knots to be that of points selected by independent Bernoulli sampling from U with success probability p. This distribution approximates that of an independent uniformly distributed set of points K_0 in D with the expected number of points, $EM_0 = Np$. It approximates the distribution of points in D generated by an n-dimensional Poisson process, conditional on M_0 points falling in D. The value of p is chosen to make operations with matrices of dimension EM_0 computationally feasible.

Given an initial knot set K_0 chosen from U as described, we use a Metropolis-Hastings algorithm to obtain successive knot sets $K_1, K_2, \ldots$. For any subset $K \subset U$ we shall use the notation K^c for the set $U \setminus K$ and $n(K)$ for the cardinality of K. The recursive construction of the knot sets can then be described as follows for $t = 1, 2, \ldots$.

(1) Draw A and B from K_{t-1} and K_{t-1}^c by Bernoulli trials with success probabilities q_1 and $q_2(t) = q_1 n(K_{t-1})/n(K_{t-1}^c)$, respectively, and define K' by $(K_{t-1} \setminus A) \cup B$. (If $n(K_{t-1}) = 0$ or $q_2(t) \geq 1$ then K' is obtained by independent Bernoulli trials with success probability p from U).

(2) Compute the acceptance ratio

$$\alpha = min\left(1, \frac{L(S_n|K')p(K')J(K' \to K_{t-1})}{L(S_n|K_{t-1})p(K_{t-1})J(K_{t-1} \to K')}\right),$$

where $L(S_n|K)$ is the conditional likelihood in (33) given the knot set K, and $p(K)$ is the prior distribution assigned for the knot set K, i.e.,

$$p(K) = p^{n(K)}(1-p)^{n(K^c)},$$

and $J(K \to K')$ is the probability of a transition from K to K', i.e.,

$$J(K \to K') = (1-q_1)^{n(K \cap K')} q_1^{n(K' \setminus K)} (1 - q_2(t))^{n(K^c \cap K'^c)} q_2(t)^{n(K'^c \setminus K^c)},$$

slightly modified if $q_2(t) \geq 1$. The other quatities $p(K'), J(K' \to K)$ are defined in the same way.

(3) Set K_t equal to K' with probability α and to K_{t-1} with probability $1 - \alpha$.

We note that p is chosen so that the expected number of initial knot points is within computational feasibility (i.e. so that Np is of the order of one hundred). The probability q_1 is chosen to be small (typically between .01 and .03) so as to give reasonable acceptance ratios and the probability $q_2(t)$ is designed to make $En(K_t)$ equal to $En(K_0) = Np$ for all t.

4.3 Kriging

Kriging, i.e. prediction of the field at points for which no observations are available, is conducted in the course of the Gibbs sampling jointly with the estimation of the parameters. Here we demonstrate kriging at a point s_0 based on the model in (32). For a given Ω at a step in the Gibbs sampling, we draw $Y = (Y_1, \ldots, Y_M)'$ from the posteriors $N(\mu_Y, \delta^2\Sigma_Y)$, where

$$\Sigma_Y^{-1} = \kappa^{-2} I_M + V_\theta' V_\theta,$$

and

$$\mu_Y = \Sigma_Y V_\theta'(S_n - Z\beta).$$

Using the sampled Y, we construct the kriged value at s_0 by

$$\hat{S}_n(s_0) = Z_0\beta + V_\theta(s_0)Y,$$

for the regressor Z_0 at s_0. Conducting the kriging in the course of the Gibbs sampling, we obtain posterior samples of the kriged value at s_0. Notice that all the procedures for the kriging as well as in the Gibbs sampling are conducted using operations with $M \times M$ matrices.

5 Empirical studies

5.1 Data generation

Simulation of the restriction of a compound Poisson CARMA random field to any bounded measurable subset $D \subset \mathbb{R}^n$ can be carried out by truncation of the infinite series in (17) as described in the following paragraph. Provided the compound Poisson sheet has finite second moments, the covariance function of the CARMA field can be computed from Theorem 3.8.

If the intensity parameter of the compound Poisson sheet is c then the number $n(D')$ of knots contained in any measurable set D' with volume $|D'|$ is simulated as a Poisson

random variable with mean $c|D'|$ and the knot locations are then simulated as $n(D')$ independent and uniformly distributed points in D'. If D' is taken to be a sufficiently large set containing D, e.g. a large hypercube containing D, then the influence of the knots outside D' will have little influence on the CARMA field restricted to D and can be neglected in the sum (17). (The required size of D' relative to that of D depends on the zero of $a(z)$ with smallest absolute value). The CARMA field restricted to D is then obtained using the sum (17) truncated to include only the simulated knots in D'. It remains only to simulate the $n(D')$ i.i.d. random variables Y_i to complete the evaluation of the simulated field at any point in D. If we take Y_i to have the normal distribution with mean zero and variance τ^2 then, in the notation of (3) and (4), $\mu = 0$ and $\sigma^2 = c\tau^2$.

Using Remark 2.6 it is also a simple matter to simulate spatial data with the Matérn covariance function,

$$\gamma_n(t) = \psi^2(||at||)^{n/2+2\nu} K_{n/2+2\nu}(||at||),\ \nu > 0,$$

using the procedure in the last paragraph with the CARMA kernel in (17) replaced by the Matérn kernel in (26), and with c and τ^2 chosen so that

$$c\tau^2 = 2\psi^2 \left(\frac{a^2}{2\pi}\right)^{n/2} \frac{\Gamma(n+2\nu)}{\Gamma(n/2+\nu)^2}.$$

5.2 Simulation studies

This section examines the empirical performance of our proposed estimation and kriging procedure using simulated compound Poisson CARMA fields on $\mathbb{R}^2$ in the model (32). In particular, we shall compare its performance with the corresponding procedure when the knots are fixed, rather than sampled.

We generated 100 sets of 1100 irregularly spaced points on the disc $D = \{(x, y) \in \mathbb{R}^2 | (x-50)^2 + (y-50)^2 < 40^2\}$. At each set of 1100 points, values of a CARMA random field were generated to represent observations of the field at those points. The 1100 data points so obtained were divided randomly into two sets of sizes 1000 and 100. The first 1000 points were used for estimation, while the remaining 100 were used for evaluating kriging mean squared errors. The values of the field at the 1100 locations were generated from (32) with zero mean function. The values of the field were simulated using the procedure described in Section 5.1 with $D' = [0, 100]^2 \supset D$. The intensity of the compound Poisson process was taken to be .02 so that the number $n(D')$ of knots uniformly and independently distributed on D' has the Poisson distribution with mean 200.

The following three CARMA kernel functions were employed in the the simulations. All the kernels were normalized to be 1 at the origin to guarantee the identifiability of τ^2, the variance of Y_j.

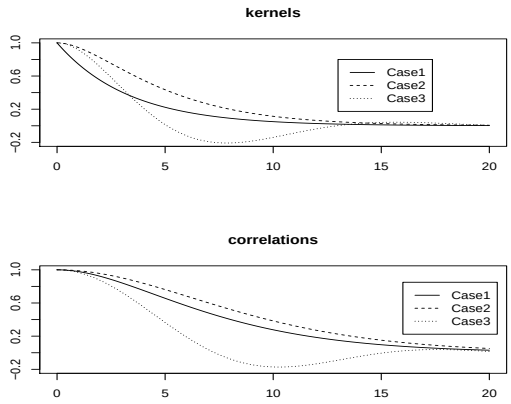

Figure 1: The kernel and resulted autocorrelation functions of the three CAR models used in the simulation studies.

Model 1. CAR(1) kernel with defining polynomial $a_*(z) = z + \alpha_1$:

$$g(s,x) = e^{-\alpha_1||s-x||}, \alpha_1 > 0.$$

Model 2. CAR(2) kernel with defining polynomial $a_*(z) = (z+\alpha_1)(z+\alpha_2)$:

$$g(s,x) = (\alpha_1 - \alpha_2)^{-1}\left(\alpha_1 e^{-\alpha_2||s-x||} - \alpha_2 e^{-\alpha_1||s-x||}\right), \alpha_1 > \alpha_2 > 0.$$

Model 3. CAR(2) kernel with defining polynomial $a_*(z) = (z+\alpha_1+i\alpha_2)(z+\alpha_1-i\alpha_2)$:

$$g(s,x) = e^{-\alpha_1||s-x||}\left\{\cos(\alpha_2||s-x||) + \frac{\alpha_1}{\alpha_2}\sin(\alpha_2||s-x||)\right\}, \alpha_1 > 0, \alpha_2 > 0.$$

The true values for the parameters were $\alpha_1 = 0.3$ in Model 1, $\alpha_1 = 0.5, \alpha_2 = 0.3$ in Model 2 and $\alpha_1 = 0.2, \alpha_2 = 0.4$ in Model 3. We used the values $\delta^2 = 1$ and $\kappa^2 = \tau^2/\delta^2 = 16$ to simulate samples from (32). The kernel functions and corresponding correlation functions in Models 1-3 are shown in Figure 1. These show that the CAR(2) kernels are smoother at the origin than the CAR(1) kernel at the origin, and that the CAR(2) kernel with complex roots can yield negative values for both the kernel and correlation functions.

We fitted the model (32) with unknown constant mean function μ to the first 1000 simulated values and evaluated the mean squared error of kriging for the remaining 100 simulated values. The prior distribution for our knot locations in D corresponded to Bernoulli trials with success probability $p = 0.1$ applied to the first 1000 data points, so that the expected initial number of knots was 100. We chose $q_1 = 0.01$ for the updating of knots in the MH algorithm with the initial knot set drawn as already described. We ran 1000 iterations of Gibbs sampling after a burn-in period of 1500 iterations. 100 posterior samples for the parameters and kriging were collected by thinning, using every 10th

	selected knots				benchmarks by fixed knots			
	α_1	α_2	κ^2	δ^2	α_1	α_2	κ^2	δ^2
Model 1	0.3	-	16	1	0.3	-	16	1
median	0.23		10.41	1.06	0.08		15.47	1.91
quart. dev.	0.01		1.34	0.05	0.02		7.75	0.15
Model 2	0.5	0.3	16	1	0.5	0.3	16	1
median	0.44	0.28	17.87	1.06	0.18	0.14	88.58	1.90
quart. dev.	0.04	0.03	2.67	0.04	0.02	0.01	34.80	0.14
Model 3	0.2	0.4	16	1	0.2	0.4	16	1
median	0.19	0.39	14.33	1.11	0.13	0.37	4.86	3.24
quart. dev.	0.01	0.01	1.79	0.04	0.01	0.01	1.63	0.38

Table 1: The median and quartile deviations of the posterior medians of our proposed estimators conducted for 100 sets of 100 samples in Models 1-3. The corresponding values with the knots fixed at a randomly chosen 100 locations are shown for comparison.

iteration. The median and quartile deviations of the posterior medians for 100 sets of 1000 simulated samples are shown in Table 1, and the mean squared errors of kriging over 100 sets of 100 samples are listed in Table 2. Estimation and kriging with the knots fixed at a randomly chosen 100 locations were conducted as benchmarks, and the medians and quartiles are listed in Tables 1-2 for comparison with our algorithm. In Table 2, as a further benchmark, we included those of simple weighted averages, whose specific form at u is

$$\left(\sum_{j=1}^{T} w_j\right)^{-1} \sum_{j=1}^{T} w_j S(s_j), \tag{34}$$

for the weight function

$$w_j = e^{-||u-s_j||^2/(bdw)^2},$$

where *bdw* denotes the bandwidth controlling the rate of decay of the weights.

We see at once that our knot specification procedure improves dramatically on the use of fixed knots for both estimation and kriging. The biases of the estimators are greatly reduced although in Model 1 there is a noticeable negative bias in the estimator of α_1. This may be due to the lack of smoothness of the CAR(1) kernel at zero. It can be reduced by increasing the value of p to increase the number of knots. In general it seems that kernels which are less smooth at the origin require more knots for accurate estimation. The improvement in kriging performance achieved by the CAR models over the weighted averages was more noticeable in Models 2 and 3 than in Model 1, which suggests that the

	CAR		weighted average			
	selected	fixed	bdw=1	2	3	4
Model 1						
median	1.38	2.05	1.85	1.45	1.50	1.69
quart. dev.	0.15	0.29	0.19	0.16	0.18	0.19
Model 2						
median	1.33	2.08	1.94	1.51	1.60	1.94
quart. dev.	0.14	0.25	0.21	0.16	0.18	0.20
Model 3						
median	1.50	3.37	2.45	2.07	2.78	3.97
quart. dev.	0.18	0.53	0.33	0.26	0.33	0.46

Table 2: The medians and quartile deviations of the mean squared errors of kriging. The medians were the medians of posterior samples and are based on 100 sets of 100 samples using each of the models 1, 2 and 3.. For comparison we also show the corresponding results for weighted averages as defined in (34) with bandwdths 1, 2, 3 and 4, and for CAR kriging with fixed knots..

weighted average procedure with suitably chosen bandwidth does a better job of capturing the covariance structure of CAR(1) data than of the data generated by higher-order CAR models.

5.3 Land price data analysis

This section demonstrates the application of the families of isotropic CAR models specified in Models 1, 2 and 3 to the analysis of Tokyo land-price data. Public land prices in Tokyo, sampled at points irregularly scattered over Tokyo's 23 wards, are used. Since 1970 the government of Japan has annually published land prices per square metre as of January 1st at hundreds of thousands of sampling points scattered irregularly all over Japan. They are evaluated using a combination of transaction records, incomes and cost accounting methods and do not therefore coincide with real market prices. They are published in March every year on the government web page in order to help in the planning of public works, inheritance tax and related economic issues.

This paper focuses on land prices in Tokyo's 23 wards collected in 2015 from

http://nlftp.mlit.go.jp/ksj/old/old_datalist.html (in Japanese).

The set of 1247 sampling points is shown in Figure 2. The price $P_t(s)$ at location s in the

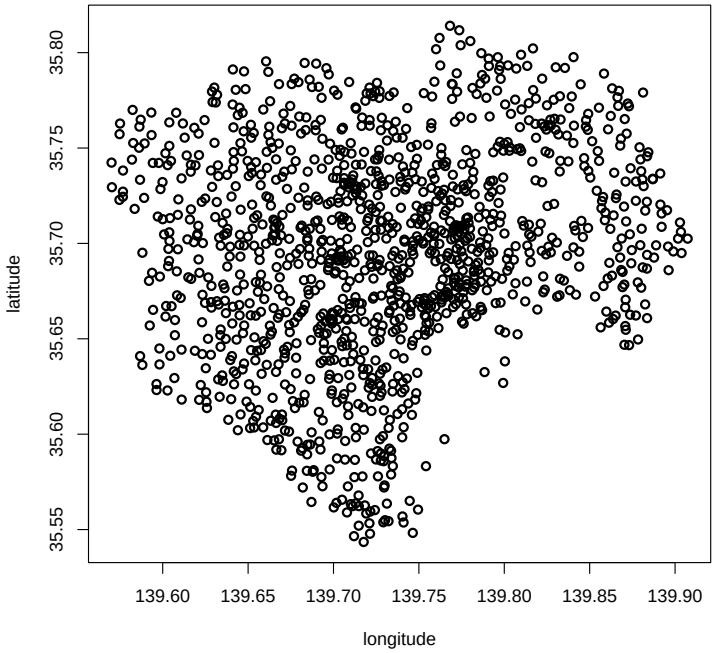

Figure 2: 1247 sampling points for public land price data in Tokyo 23 wards collected in 2015.

year $t = 2015$ was transformed to a log return $r_t(s)$, given by

$$r_t(s) = \log P_t(s) - \log P_{t-1}(s), t = 2015.$$

The transformed great circle distance (Gneiting, 1999) was used to measure the distance $d(s, s')$ between the locations s and s'. Thus

$$d(s, s') = 2r \sin(\Delta\phi/2),$$

where r is radius of the earth in kilometers and $\Delta\phi \in [0, \pi]$ is the central angle between s and s'. We divided the 1247 locations randomly into two groups. The first group of 1147 was used for model fitting and the second group of 100 was used for evaluating kriging mean squared errors.

We fitted the model (32) with constant mean function μ and CAR models 1, 2 and 3 to the returns at the first 1147 locations and evaluated the mean squared errors of kriging at the remaining 100. We chose $p = 100/1147$ to give an expected number of knots equal to 100 and $q = 0.01$ in the knots updating step. We ran 10000 iterations of the Gibbs sampling after a burn-in period of 10000 iterations. 1000 posterior samples were collected and these were thinned by using every 10th iteration.

	selected knots						benchmarks by fixed-knots				
	α_1	α_2	κ^2	δ^2	M	logL	α_1	α_2	κ^2	δ^2	logL
Model 1											
median	1.03	-	22.14	0.49	113	209.7	0.22	-	2.25	1.07	-117.2
quart. dev.	0.04	-	3.64	0.01	4.5	7.7	0.02	-	0.54	0.01	0.6
Model 2											
median	15.10	1.02	18.41	0.54	101	168.3	1.51	0.29	2.05	1.04	-112.2
quart. dev.	1.41	0.05	3.89	0.02	5.5	12.5	0.23	0.04	0.45	0.01	0.7
Model 3											
median	2.11	0.10	9.86	0.63	105	97.8	0.55	0.08	1.83	1.04	-113.4
quart. dev.	0.06	0.07	1.24	0.01	3.5	7.1	0.04	0.05	0.38	0.01	0.8

Table 3: CAR fitting by Models 1-3 for 1147 samples of log returns of land prices in Tokyo. The median and the quartile deviation for the posterior samples were evaluated in comparison with those of the cases of fixed knots with 100 randomly chosen locations.

In Table 3, we show the medians and quartile deviations of posterior samples with their log likelihoods as given in (33). We conducted the fitting also with knots fixed at 100 randomly chosen locations and listed the results in Table 3 as benchmarks. In Table 4, we show the mean squared errors of kriging using the posterior medians of the kriging samples as kriged values. The performance of these kriged values is compared with that of the weighted average (34) and the corresponding kriged values obtained using fixed knots.

We find that in terms of log likelihood our knot selection procedure improves significantly on the fixed knot procedure and the CAR(1) model outperforms the CAR(2) models. The CAR(1) model with our knot selection procedure has the best kriging mean squared error of all considered. It is interesting that the improvement in kriging performance of the knot selection procedure over the fixed knot procedure is not nearly as pronounced as in the simulation study.

The kernel and autocorrelation functions of the best fitting CAR(1) are shown in Figure 3 and the corresponding smoothed values, obtained by kriging over the entire sampling region of Tokyo's 23 wards are shown in Figure 4. According to Figure 3 the log returns of land prices have autocorrelations smaller than .05 for lags greater than 6 kilometres. Figure 4 clearly demonstrates the interesting shape of the returns surface, with the highest returns in the centre of Tokyo, gradually decreasing as we move towards the remote suburbs.

	selected knots			fixed knots			weighted average			
	Model1	Model2	Model3	Model1	Model2	Model3	bdw=1	2	3	4
MSE	0.93	0.96	1.10	1.02	1.03	1.05	1.05	1.03	1.08	1.17

Table 4: Kriging mean squared errors for randomly chosen 100 samples in log returns of land price data in Tokyo, where kriging was conducted as the median of posterior samples. The weighted averages with bandwidth of 1, 2, 3 and 4 and kriging with fixed knots were conducted as benchmarks.

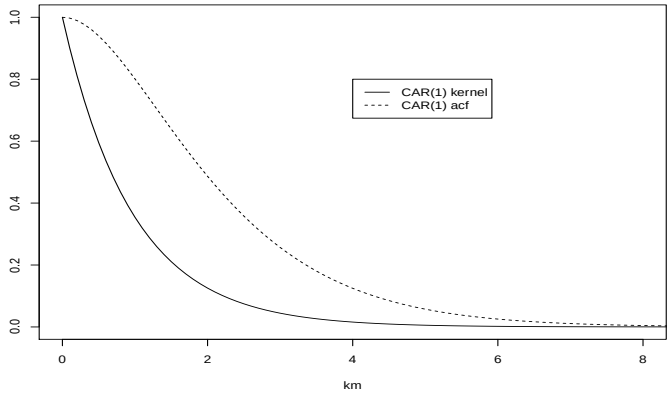

Figure 3: The identified kernel and autocorrelation functions by Model 1 for log returns of land prices in Tokyo 23 wards.

6 Discussion

This paper proposes the use of isotropic CARMA random fields driven by Lévy sheets and specified by a kernel function defined in terms of autoregressive and moving average polynomials of orders p and q respectively where $0 \leq q < p$. When the Lévy sheet has finite second moments this class generates a rich source of isotropic covariance functions on $\mathbb{R}^n$ with spectral densities which can be written explicitly. The covariance functions are not necessarily non-negative or monotone. The covariance function of a CAR(1) field has a simple explicit expression which belongs to the Matérn class with smoothness parameter 2. The covariances of more general CARMA fields can be expressed as modified Hankel transforms which can be explicitly evaluated when $n = 1$ or $n = 3$.

Compound Poisson CARMA random fields are particularly useful as their restriction to any bounded measurable subset of $\mathbb{R}^n$ can be approximated by a finite sum of terms involving no matrix operations. Moreover estimation and kriging can be jointly conducted by Gibbs sampling requiring $O(M^3)$ operations where M is the number of terms used in the approximation to the random field. The Metropolis-Hastings algorithm can be used

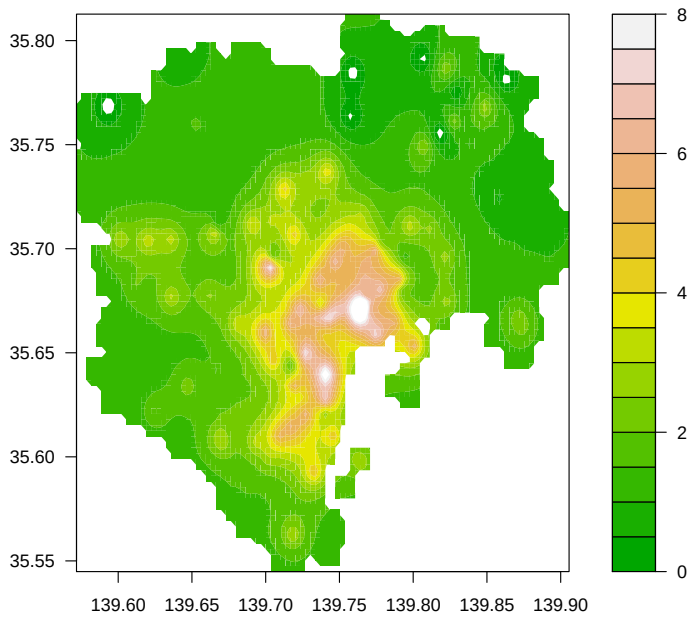

Figure 4: The smoothed figure of log returns of land price in Tokyo 23 wards by Model 1.

to specify the knot locations in the Gibbs sampling and the procedure works well in the simulated examples and in terms of kriging mean squared error in the Tokyo land-price example.

References

[1] Abramowitz, M. and Stegun, I. A. (2012). *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*. Dover Publications, New York.

[2] Brockwell, P.J. and Lindner, A. (2009). Existence and uniqueness of stationary Lévy-driven CARMA processes, *Stoch.Proc.Appl.* 119, 2660–2681.

[3] Brockwell, P.J. (2014). Recent results in the theory and applications of CARMA processes, *Ann. Inst. Statistical Mathematics* 66, 647–685.

[4] Calder, C.A. and Cressie, N. (2007). Some topics in convolution-based spatial modeling, In: *Proceedings of the 56th Session of the International Statistical Institute*, Lisbon, International Statistical Institute. Voorburg, the Netherlands.

[5] Fuentes, M. (2002). Spectral methods for nonstationary spatial processes. *Biometrika*, 89, 197-210.

[6] Gelfand, A. E., Banerjee, S. and Finley, A. O. (2012). Spatial Design for Knot Selection in Knot-Based Dimension Reduction Models, in Spatio-Temporal Design: Advances in Efficient Data Acquisition (eds J. Mateu and W. G. Müller), John Wiley & Sons, Ltd, Chichester, UK.

[7] Gneiting, T. (1999). Correlation functions for atmospheric data analysis, *Quarterly J. Roy. Meteorological Soc.* 125, 2449–2464.

[8] Grafakos, L. and Teschl, G. (2013). On Fourier transforms of radial functions and distributions, *J. Fourier Anal. Appl.* 19, 167–179.

[9] Harville, D. A. (1997). *Matrix Algebra from a Statistician's Perspective*. Springer, New York.

[10] Higdon, D. (1998). A process-convolution approach to modeling temperatures in the North Atlantic Ocean. *Environmental and Ecological Statistics*, 5, 173-190.

[11] Higdon, D. (2002). Space and space-time modeling using process convolutions. *Quantitative Methods for Current Environmental Issues.*, 37-56.

[12] Khoshnevisan, D. and Nualart, E. (2008). Level sets of the stochastic wave equation driven by a symmetric Lévy noise,*Bernoulli* 14, 899–925.

[13] Knockaert, L. (2000). Fast Hankel transform by fast sine and cosine transforms: the Mellin connection,*IEEE Transactions on Signal processing* 48, 1695–1701.

[14] Majumdar, A., Paul, D. and Bautista, D. (2010). A generalized convolution model for multivariate nonstationary spatial processes. *Statistica Sinica*, 20, 675-695.

[15] Matérn, B. (1960). *Spatial Variation*. Meddelanden frán Statens Skogsforskningsinstitut, 49, Almaenna Foerlaget, Stockholm. Springer-Verlag, Berlin.

[16] Nowak, A. and Stempak, K. (2014). A note on recent papers by Grafakos and Teschl, and Estrada, *J. Fourier Anal. Appl.* 20, 1141–1144.

[17] Paciorek, C.J. and Schervish, M. J. (2006). Spatial modeling using a new class of nonstationary covariance functions. *Environmetrics*, 17, 483-506.

[18] Piessens, R. (2010). The Hankel transform. Chapter 9 of *The Transforms and Applications Handbook: Third Edition*, ed. Alexander D. Poularikas, CRC Press, Boca Raton.

[19] Sampson, P.D. (2010). Cnstructions for nonstationary spatial processes. *Handbook of Spatial Statistics*. Edited by Gelfand, A.E. Diffle, P., Fuentes, M. and Guttorp, P., 119-130. CRC Press, Boca Raton.

[20] Schnurr, A. and Woerner, J. H. C. (2011). Well-balanced Lévy-driven Ornstein-Uhlenbeck processes, *Statistics and Risk Modelling* 4, 343–359.

[21] Stein, M. L. (1999). *Interpolation of Spatial data: Some Theory for Kriging*. Springer, New York.

[22] Zhang, B., Sang, H. and Huang, J.Z. (2015). Full-scale approximations of spatio-temporal covariance models for large datasets. *Statistica. Sinica* 25, 99–114.